

What is claimed is:

1. A method for transferring data through a packet switch while providing differentiated Quality-of-Service (QoS) guarantees to respective traffic flows, comprising the steps of:
- storing received data packets associated with said traffic flows in a respective input buffer;
  - grouping said traffic flows;
  - selecting stored data packets for transmission to a switch fabric according to a first plurality of schedulers;
  - assigning bandwidth to said selected data packets according to a second plurality of schedulers;
  - storing said transmitted data packets in a plurality of output buffers in said switch fabric; and
  - adapting the operation of at least one of said first and second schedulers in response to a determination that a utilization level of any output buffer has exceeded a threshold parameter.
2. The method of claim 1, further comprising the step of:
- choosing the data packets to be transmitted out of said plurality of output buffers according to a third plurality of schedulers.
3. The method of claim 1, wherein said step of grouping comprises the step of grouping said traffic flows according to traffic flow destination and QoS requirements.
4. The method of claim 3, wherein said destination comprises an output of an egress port card coupled to said switch fabric.
5. The method of claim 1, wherein each scheduler in said first plurality of schedulers is a class-specific per-flow scheduler.

002290-05266560

6. The method of claim 1, wherein each scheduler in said second plurality of schedulers is a per-QoS-class port scheduler.

7. The method of claim 3, wherein each of said second plurality of  
5 schedulers comprises at least one of:

a Guaranteed Bandwidth Scheduler (GBS), for selecting said groups of configured traffic flows for transmission of data packets based on minimum bandwidth guarantees for the QoS channels;

an Excess Bandwidth Scheduler (EBS), for redistributing the unused  
10 GBS bandwidth to QoS channels waiting for transmission of data packets to said respective plurality of output buffers; and

a priority scheduler for assigning priority to the GBS over the EBS for the transmission of data packets.

15 8. The method of claim 2, wherein each of said third plurality of schedulers comprises at least one of:

a Guaranteed Bandwidth Scheduler (GBS), for selecting said groups of configured traffic flows for transmission of data packets based on minimum bandwidth guarantees for the QoS channels;

20 an Excess Bandwidth Scheduler (EBS), for redistributing the unused GBS bandwidth to QoS channels waiting for transmission of data packets from said respective plurality of output buffers; and

a priority scheduler for assigning priority to the GBS over the EBS for the transmission of data packets.

25

9. The method of claim 1, wherein if said threshold parameter has been exceeded in any of said plurality of output buffers, said method further comprises the steps of:

applying selective backpressure to respective ingress port cards to  
30 prevent further transmission of data packets to any of said respective plurality of output buffers for which the threshold parameter has been

002290 05266560

exceeded, said plurality of output buffers for which the respective threshold parameter has been exceeded being congested;

determining whether said respective threshold level of each of said congested plurality of output buffers continues to be exceeded; and

5 resuming transmission to any congested respective plurality of output buffers when each respective said threshold parameter is not exceeded.

10. The method of claim 1, wherein:

each scheduler of said first plurality of schedulers resides in ingress  
10 port cards;

each scheduler of said second plurality of schedulers resides in said ingress port cards;

each scheduler of said third plurality of schedulers resides in said switch fabric;

15 said plurality of input buffers reside in said ingress port cards; and said plurality of output buffers reside in said switch fabric.

11. The method of claim 10, wherein said respective plurality of input buffers' buffer capacity exceeds said respective plurality of output buffers' buffer capacity.  
20

12. The method of claim 1, wherein said switch fabric comprises at least one module having a respective plurality of output buffers.

25 13. The method of claim 1, wherein each of said first plurality of schedulers is associated with a respective QoS class.

14. The method of claim 1, wherein each of said plurality of output buffers supports aggregated traffic flows of a defined QoS class.

30

15. The method of claim 1, wherein each one of said third plurality of schedulers is associated with a respective output of said switch fabric.

002290-05266560

16. A method for transferring data packets through a packet switch while providing differentiated Quality-of-Service (QoS) guarantees to respective traffic flows, comprising the steps of:

5 storing incoming data packets associated with traffic flows in a plurality of input buffers;

grouping said traffic flows by QoS classes, each of said QoS classes having a priority level;

10 transmitting data packets to a respective plurality of output buffers according to QoS class priority level;

determining whether the utilization level of any of a plurality of said output buffers has exceeded a threshold parameter; and

halting transmissions to any of said plurality of output buffers utilized beyond said threshold parameter.

15

17. The method of claim 16, further comprising the step of:

resuming transmission to any of said output buffers having utilization levels below said threshold parameter.

20 18. The method of claim 16, wherein said priority level is determined for each traffic flow by a scheduler,

said scheduler comprising at least one of:

25 a Guaranteed Bandwidth Scheduler (GBS), for selecting said groups of configured traffic flows for transmission of data packets based on minimum bandwidth guarantees for the QoS classes;

an Excess Bandwidth Scheduler (EBS), for redistributing the unused GBS bandwidth to QoS classes waiting for transmission of data packets to said respective plurality of output buffers; and

30 a priority scheduler for assigning priority to the GBS over the EBS for the transmission of data packets.

002290 0526560

19. The method of claim 16, wherein if said threshold parameter has been exceeded in any of said plurality of output buffers, said method further comprises the steps of:

- 5 applying selective backpressure to respective ingress port cards to prevent further transmission of data packets to any of said respective plurality of output buffers for which the threshold parameter has been exceeded, said plurality of output buffers for which the respective threshold parameter has been exceeded being congested;
- 10 determining whether said respective threshold level of each of said congested plurality of output buffers continues to be exceeded; and resuming transmission to any congested respective plurality of output buffers when each respective said threshold parameter is not exceeded.

20. The method of claim 16, each scheduler of said first plurality of
- 15 schedulers resides in ingress port cards;
- each scheduler of said second plurality of schedulers resides in said ingress port cards;
  - each scheduler of said third plurality of schedulers resides in said switch fabric;
  - 20 said plurality of input buffers reside in said ingress port cards; and said plurality of output buffers reside in said switch fabric.

21. An apparatus for transferring data packets through a packet switch while providing differentiated Quality-of-Service (QoS) guarantees,
- 25 comprising:
- a first plurality of schedulers for selecting traffic flows arranged in groups;
  - a second plurality of schedulers coupled to said first plurality of schedulers for assigning bandwidth to said selected groups of traffic flows;
  - 30 a plurality of input buffers, coupled to said first and second pluralities of schedulers, for holding data packets associated with said grouped traffic flows;

002290-052656

a third plurality of schedulers, for selecting data packets for transmission to respective output ports;

a plurality of output buffers, coupled to said third plurality of schedulers, for holding data packets before transmission to said respective  
5 output ports; and

a backpressure-signal circuit connected between each of said respective plurality of output buffers and each of said respective plurality of input buffers, transmitting a stop-transmission signal to each of said respective plurality of input buffers when a threshold parameter in any of  
10 said respective plurality of output buffers is exceeded.

22. The apparatus of claim 21, wherein:

said first plurality of schedulers resides in ingress port cards;  
said second plurality of schedulers resides in said ingress port cards;  
15 said plurality of input buffers reside in said ingress port cards;  
said third plurality of schedulers resides in a switch fabric;  
said output ports reside in said switch fabric;  
said plurality of output buffers reside in said switch fabric; and  
said ingress port cards are coupled to said switch fabric.

20

23. The apparatus of claim 21, wherein each of said second plurality of schedulers comprises at least one of:

a Guaranteed Bandwidth Scheduler (GBS), for selecting said groups of configured flows for transmission of data packets based on minimum  
25 bandwidth guarantees for QoS classes;

an Excess Bandwidth Scheduler (EBS), for redistributing unused GBS bandwidth to QoS classes waiting for transmission of data packets to said respective plurality of output buffers; and

a priority scheduler for assigning priority to the GBS over the EBS for  
30 the transmission of data packets.

002290 0526560

24. The apparatus of claim 21, wherein each of said third plurality of schedulers comprises at least one of:

a Guaranteed Bandwidth Scheduler (GBS), for selecting data packets for transmission based on minimum bandwidth guarantees for QoS

5 channels;

an Excess Bandwidth Scheduler (EBS), for redistributing unused GBS bandwidth to QoS channels waiting for transmission of data packets from said respective plurality of output buffers; and

a priority scheduler for assigning priority to the GBS over the EBS for  
10 the transmission of data packets.

25. The apparatus of claim 21, wherein said respective plurality of input buffers' buffer capacity exceeds said respective plurality of output buffers' buffer capacity.

15

26. The apparatus of claim 22, wherein said switch fabric comprises a single module with a plurality of output buffers.

27. The apparatus of claim 22, wherein said switch fabric comprises  
20 multiple modules having a plurality of output buffers.

28. The apparatus of claim 21, wherein each one of said first plurality of schedulers is associated with a respective QoS class of said configured flows.

25 29. The apparatus of claim 21, wherein said plurality of output buffers are coupled to traffic aggregates called QoS channels.

30. The apparatus of claim 21, wherein each one of said third plurality of schedulers is associated with a respective output of said switch fabric.

002290" 05266560